

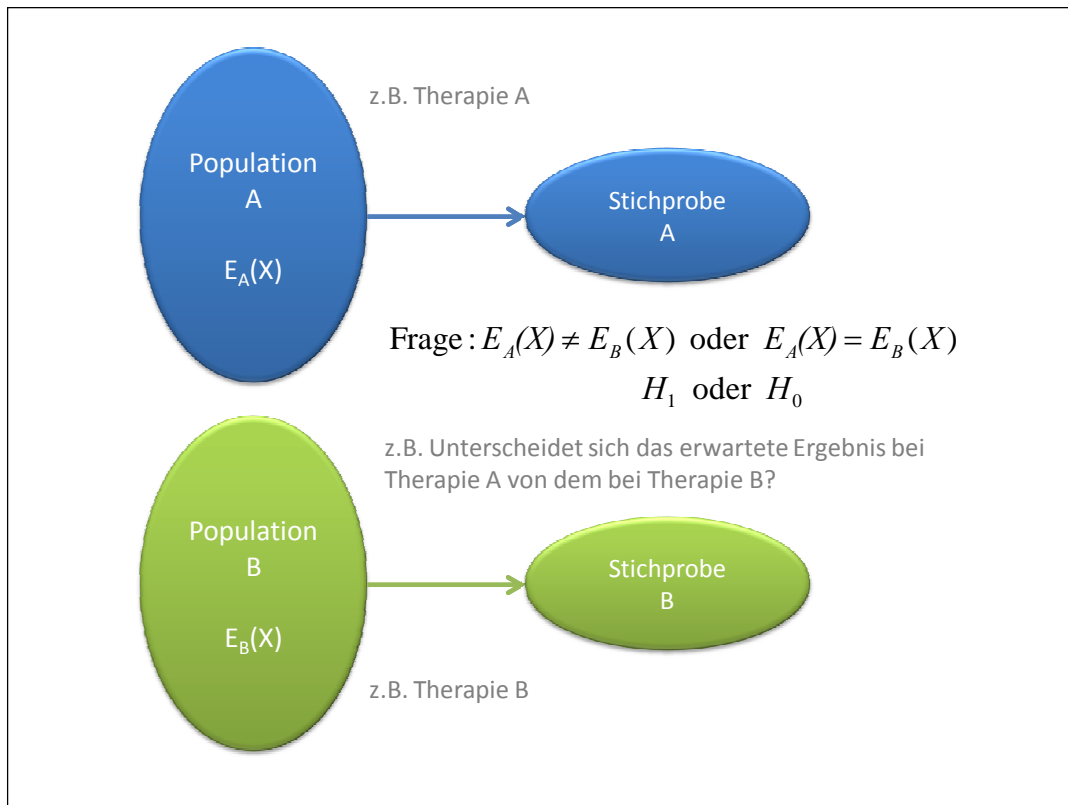
Statistik

Teil 6 - Statistisches Testen

Ulrich Schrader
<http://info.ulrich-schrader.de>

Statistisches Testen

- Situation
 - 2 Populationen aus denen jeweils eine Stichprobe gezogen wird. Ein Merkmal X wird in beiden Stichproben beobachtet.
- Frage
 - Unterscheidet sich der Erwartungswert des Merkmals in den beiden Populationen?
- 2 Hypothesen
 - "Die Erwartungswerte unterscheiden sich nicht in den beiden Populationen." - **Nullhypothese H_0**
 - "Die Erwartungswerte unterscheiden sich in den beiden Populationen." - **Alternativhypothese H_1**



Entscheidungsmatrix

| Entscheidung für | Es liegt tatsächlich vor | |
|-----------------------------|--|---|
| | H_0 – „kein Unterschied“ | H_1 – „Unterschied“ |
| H_0 „kein Unterschied“ | richtige Entscheidung (1 - α) | falsch negative Entscheidung β – „Fehler 2. Art“ |
| H_1 „Unterschied“ | falsch positive Entscheidung α – „Fehler 1. Art“ | richtige Entscheidung (1 - β) |

Ein *statistischer Test* ist eine Entscheidungsregel, welche die Wahrscheinlichkeit für einen Fehler 1. Art kleiner als einen vorgegebenen Wert α hält und die Wahrscheinlichkeit β für einen Fehler 2. Art möglichst gering hält.

- $1 - \alpha$: Signifikanzniveau
- α : Irrtumswahrscheinlichkeit
- $1 - \beta$: Güte/Mächtigkeit des Tests

Vorgehen beim statistischen Testen

1. Schritt

- Formulierung der Hypothesen
 - Alternativhypothese und Nullhypothese
- Festlegung der noch zulässigen Irrtumswahrscheinlichkeit α (Fehler 1. Art)
 - Wahrscheinlichkeit eines Fehlers im Fall der Entscheidung für die Alternativhypothese H_1)
 - Bestimmt sich aus den möglichen Folgen einer Fehlentscheidung
- Auswahl eines geeigneten Tests
 - Fehler 2. Art β möglichst klein bei vorgegebenem Signifikanzniveau $(1 - \alpha)$
(Möglichst mächtiger Test)

Vorgehen beim statistischen Testen

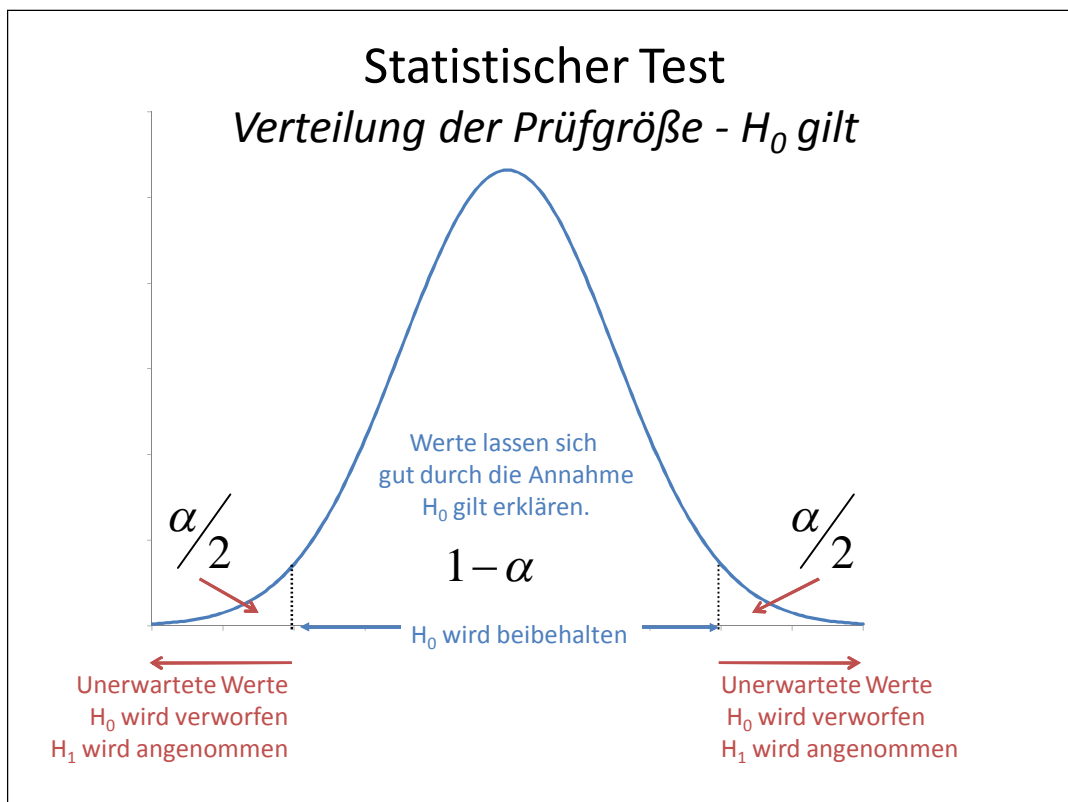
2. Schritt

- Annahme (willkürlich)
 - H_0 gilt (kein Unterschied)
(Konservativ: Die neue Therapie macht keinen Unterschied)
- Berechnung einer Prüfgröße, deren Verteilung bekannt ist, wenn die Annahme H_0 erfüllt ist, aus den Werten der Stichproben.

Vorgehen beim statistischen Testen

3. Schritt

- Verwerfen der ursprünglichen (willkürlichen) Annahme, H_0 soll gelten, wenn der Wert für die Prüfgröße gemäß der bekannten Verteilung nicht erwartet wird. Dann Alternativhypothese H_1 annehmen.
- Sonst H_0 beibehalten. Konservativer Ansatz.





p-Wert

- Oft als Ergebnis eines statistischen Tests angegeben.
- Geringste Irrtumswahrscheinlichkeit α , deren Festlegung mit den beobachteten Stichprobenwerten gerade noch zur Verwerfung der Nullhypothese geführt hätte.

Beispiel

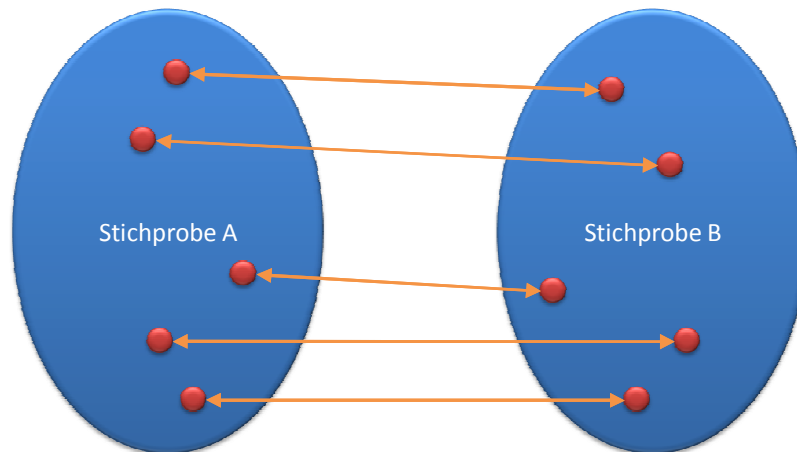
IM Rosen, et al. : Prenatal breastfeeding education and breastfeeding outcomes. *MCN. The American Journal of Maternal Child Nursing*, Vol. 33, No. 5. (008), pp. 315-319.

- **PURPOSE: To examine the impact of various breastfeeding outcomes of three cohorts receiving different methods of prenatal breastfeeding education.**
METHODS: Retrospective cohort design with patients attending a breastfeeding education class at an Army medical center. Controls were matched for sponsor rank, marital status, and smoking status. One hundred ninety-four mothers who expressed intent to breastfeed received breastfeeding education as follows: (a) a class that used video demonstration and group teaching by a lactation consultant, (b) a new mothers' support group with one-on-one teaching prenatally and weekly meetings postpartum, taught by a lactation consultant and a pediatrician, and (c) a control group educated at prenatal visits only. Descriptive statistics, chi-square, ANOVA, **unpaired t test**, and logistic regression were used to analyze the data.
- **RESULTS: Women who attended prenatal breastfeeding classes had significantly increased breastfeeding at 6 months when compared to controls ($p = .01$). There was no significant difference in rates between types of classes offered ($p = .45$).**
- **CLINICAL IMPLICATIONS:** Prenatal breastfeeding education can influence the amount of time women breastfeed. All providers of prenatal care should consider offering such classes in order to improve breastfeeding rates.

Tests mit unabhängigen/paarigen Stichproben

- **Unabhängige Stichproben**
 - Beobachtungseinheiten der einen Stichprobe haben keine Beziehung zu Beobachtungseinheiten der anderen Stichprobe
 - Indiz: Stichprobenanzahlen unterschiedlich
- **Paarige Stichproben**
 - Jede Beobachtungseinheit der einen Stichprobe hat genau eine korrespondierende Beobachtungseinheit in der anderen Stichprobe.
 - Typisch: Vorher/Nachher
Mutter in der einen, Kind in der anderen Stichprobe
Zwillinge
 - Notwendig: Stichprobenanzahlen gleich

Paarige Stichproben



Übersicht über wichtige statistische Testverfahren

| Merkmal | Prüfgröße | Unabhängige SP | Paarige SP |
|---|-------------------------|----------------------------------|-----------------------------------|
| Quantitativ, normalverteilt, ähnl. Streuung | t-verteilt | Student's t-Test für unabh. SP | Student's t-Test für paarige SP |
| Quantitativ, beliebig verteilt | standard-normalverteilt | Wilcoxon Rang-Test für unabh. SP | Wilcoxon Rang-Test für paarige SP |
| Ordinal | standard-normalverteilt | Wilcoxon Rang-Test für unabh. SP | Wilcoxon Rang-Test für paarige SP |
| Nominal (dichotom, zweiwertig) Häufigkeiten | χ^2 -verteilt | Chi-Quadrat-Vierfeldertest | Vorzeichentest |

SP : Stichprobe(n)

Beispiel

MY Chang, CH Chen, KF Huang: Effects of music therapy on psychological health of women during pregnancy. Journal of clinical nursing, Vol. 17, No. 19. (October 2008), pp. 2580-2587.

- AIMS AND OBJECTIVES: The purpose of this study was to examine the effects of music therapy on stress, anxiety and depression in Taiwanese pregnant women.
- BACKGROUND: The value of music therapy is slowly being realized by nurses in various clinical areas, including obstetrics. Previous studies have demonstrated a high prevalence of psychological stress during pregnancy. [...]
- DESIGN: A randomized experimental study design was developed and implemented.
- METHODS: **Two hundred and thirty-six pregnant women were randomly assigned to music therapy (n = 116) and control (n = 120) groups.** The music therapy group received two weeks of music intervention. The control group received only general prenatal care. Psychological health was assessed using three self-report measures: Perceived Stress Scale (PSS), State Scale of the State-Trait Anxiety Inventory (S-STAI) and Edinburgh Postnatal Depression Scale (EPDS). RESULTS: **In a paired t-test, the music therapy group showed significant decrease in PSS, S-STAI and EPDS after two weeks. The control group only showed a significant decrease in PSS after two weeks. This decrease was not as substantial as in the experimental group.** [...]
- CONCLUSIONS: This controlled trial provides preliminary evidence that two-week music therapy during pregnancy provides quantifiable psychological benefits.
- RELEVANCE TO CLINICAL PRACTICE: The findings can be used to encourage pregnant women to use this cost-effective method of music in their daily life to reduce their stress, anxiety and depression. Further research is needed to test the long-term benefits.

Beispiel

CL Dennis, et al.: Effect of peer support on prevention of postnatal depression among high risk women: multisite randomised controlled trial. BMJ (Clinical research ed.), Vol. 338 (2009)

OBJECTIVE: To evaluate the effectiveness of telephone based peer support in the prevention of postnatal depression.

DESIGN: Multisite randomised controlled trial. SETTING: Seven health regions across Ontario, Canada.

PARTICIPANTS: 701 women in the first two weeks postpartum identified as high risk for postnatal depression with the Edinburgh postnatal depression scale and randomised with an internet based randomisation service. I

INTERVENTION: Proactive individualised telephone based peer (mother to mother) support, initiated within 48-72 hours of randomisation, provided by a volunteer recruited from the community who had previously experienced and recovered from self reported postnatal depression and attended a four hour training session.

MAIN OUTCOME MEASURES: Edinburgh postnatal depression scale, structured clinical interview-depression, state-trait anxiety inventory, UCLA loneliness scale, and use of health services.

RESULTS: After web based screening of 21 470 women, 701 (72%) eligible mothers were recruited. A blinded research nurse followed up more than 85% by telephone, including 613 at 12 weeks and 600 at 24 weeks postpartum. At 12 weeks, 14% (40/297) of women in the intervention group and 25% (78/315) in the control group had an Edinburgh postnatal depression scale score >12 (chi(2)=12.5, P<0.001; number need to treat 8.8, 95% confidence interval 5.9 to 19.6; relative risk reduction 0.46, 95% confidence interval 0.24 to 0.62). [...] Of the 221 women in the intervention group who received and evaluated their experience of peer support, over 80% were satisfied and would recommend this support to a friend.

CONCLUSION: Telephone based peer support can be effective in preventing postnatal depression among women at high risk.

